Gu Test: A Progressive Measurement Of Generic Artificial Intelligence

Author: Lifan Gu

According to news, in 2015 a blind man was allowed to take a driverless car alone, before the accident on 02/14/2016 [1]. Although the damage of this accident is minor, wrong judgement of driverless cars is very dangerous potentially. The simulations and empirical tests in the past could not prove the good judgement in future.

According to the Waymo Safety Report 2017, the driverless car is highly wired with pre-information [2], making it very difficult to adapt to future mode evolution, especially when the mode evolution is not stable. The probability of mistakes in future could change significantly with mode evolution. Actually nothing in this safety report could test level-5 driverless cars.

The AlphaGo Zero paper on Nature claimed a superhuman performance [3]. However the paper did not provide any evidence for this claim. Superhuman is a concept related to generic human. AlphaGo Zero defeating AlphoGo Master is not an evidence. AlphoGo Master defeating some humans in some games is not an evidence, either. A scientific experiment with strictly controlled conditions could falsify the claim.

So the current testing of AI compared with humans is highly problematic. The 3rd edition of a mainstream AI textbook Artificial Intelligence: A Modern Approach cannot address these important issues. Actually it took Aristotle philosophy as the foundation by mistake.

Turing Test is invalid, not only it is subjective, but also the language complexity is much less than the human intelligence complexity.

So instead, I design Gu Test, a progressive measurement of generic artificial intelligence by their falsifiability.

Computers, including quantum computers, have systematic problems with high-order logics. Gödel theorems suggest mathematics can not be used to judge the correctness of sciences. There is limitation of Turing Machine. So universal approximation does not exist on Turing Machine, and Technological Singularity is baseless.

If the mode evolution is not stable, not only the judgement based on intuition could be wrong, the statistical results, the deep-learning of empirical data, and other AI technologies also could be severely misleading. A theoretic study of human intelligence is needed, to cross languages, philosophies, mathematics, and sciences. etc.

Based on my studies, I design certain test programs, to test some artificial intelligence systems, and also to test my theories:

1. A 4-dimension experiment space for the strongest Computer Go system, to test its intelligence on Go games, especially for AlphaGo Zero's superhuman claim due to the problems in deep-learning, reinforcement, etc.

Since there is only one opportunity to gather certain experiment results before the computer Go system could be adjusted by humans, the first round experiment should be done on the strongest Computer Go system.

2. A progressive test scheme for natural language processing (NLP) system, including high-order logics, etc.

.....

The studies could be extended to other kinds of AI systems, and other aspects of human intelligence in future.

However, life-threatening situations happened to me again and again. I cannot do further researches unless in safe personally and economically. Actually, some health degrading could be irreversible, so I may not be able to do further research except for the experiment schemes already designed.

[1] https://www.washingtonpost.com/local/trafficandcommuting/blind-man-sets-out-alone-in-googles-driverless-car/2016/12/13/f523ef42-c13d-11e6-8422-eac61c0ef74d_story.html

https://www.marketwatch.com/story/google-says-driverless-cars-are-ready-to-make-money-but-we-wont-know-if-they-do-2016-12-13

[2] https://waymo.com/safety/

[3] https://www.nature.com/articles/nature24270